

# PERSIAN LOAN WORDS IN 18<sup>TH</sup> CENTURY YENISEIC AND THE PROBLEM OF LINGUISTIC AREAS IN SIBERIA

Marek Stachowski (Kraków)

The oldest materials of the Yeniseic languages date from the 18<sup>th</sup> century and are more often than not relatively scant word lists. Their publication by E. Helimski (1986), afterwards included into VWJ, was an important contribution to Yeniseiology.

It is my aim here to see how many Persian loanwords in the Yeniseic languages can be traced as far back as possible, i.e. practically into the 18<sup>th</sup> century and to establish to what extent their existence can influence our understanding of areal processes in the Siberian linguistics.

In what follows, one will find some doubts and suggestions, rather than ultimate solutions and definite answers. Moreover, I can mainly rely on languages which became extinct (and were attested) in the 18<sup>th</sup> century, i.e. on Arinian, Assanian, Pimpokolian, Koibal (in those days a Yeniseic idiom<sup>1</sup>) and Imbatian<sup>2</sup>, since Ket, Kott and Yug, which are principally better attested, seem to record the 18<sup>th</sup> century Persian loan words but sporadically. That is why this modest article can only be viewed as the first attempt at studying Persian loan words in the Yeniseic languages. Consequently, remarks concerning the areal groupings of Siberian languages and gleaned from the analysis of the lexical material presented in the first part of the study are of a tentative character, too.

## 1.

The borrowing channels of three words can be relatively easily established:

*ajna* (Arin.) 'devil' (VWJ I 21) < Turkic, e.g. Shor, Khak., Oir. *ajna* id. << Old Pers. *hajnā*+ id. (Stachowski 1996: 102). – Unfortunately, the Old Persian etymon, first established by B. Munkácsi (1900) as the source of Ostyak *xeina* 'Führer der unterirdischen bösen Heerscharen' (ibid.) cannot be used as an argument in the discussion of the chronology of the borrowing process as it only signals the ultimate source (not the direct one) of the Turkic words. – The direct borrowing of the Arinian word from Old Persian cannot be excluded on a purely phonetic basis; however, it requires an historical scenario which makes a more or less direct contact possible.<sup>3</sup> On the other

<sup>1</sup> Cf. Joki 1952: "Einige frühere Forscher hatten die Koibalen und Kotten für eine und dieselbe ethnisch-linguistische Einheit gehalten [...]."

<sup>2</sup> Cf. VWJ I 9: "imb. = imbazkisch [ketisch 18.Jh.]".

<sup>3</sup> This aspect of our Yeniseiological knowledge still may not be viewed as perfectly understood; nevertheless, newer research points to a possible necessity of restructuring our historical conception of the role and geographical range of the Yeniseic languages, see e.g. the introductory parts in Vovin 2000, 2003. It seems also very tempting to examine once again E. G. Pulleyblank's concept, who tried to connect the name *Cāč*, being an old name of Tashkent, with Kott *šīs* (not *šis*, against Aalto 1987) = Ket *tyʔś* and other Yeniseic names for 'stone' (cf. Turkic *taš-kent*, lit. 'stone-city'). Pentti

hand, no phonetic or morphological feature of the word suggests another channel than the Turkic one.

*molát* (Arin.) ~ *mōlat* (Koib.) ~ *balat* (Ass.) ~ *bolat* (Ket: Castrén) 'steel' (VWJ I 139) = Shor (ŠS), Khak. *molat*, Tof. *bolat*, Oir. *bolot* ~ *polot* id. = Mo. *bolot* ~ *bolod* ~ *bolad* ~ *bolud* id. (Joki 1952: 96) = Mator *bolat* id. (Helimski 1997, Nr. 136) < Pers. *pūlād* id. (Stachowski 1993: 251; 2002: 8). – The *m*- variants point unequivocally to the Turkic mediation, whereas Ass. *balat* and Kott *bolat* may reflect Turkic and Mator *bolat* or Mo. *bolad*.

*nañ* ~ *nān* ~ *nān* (Imb.) 'bread' [= modern Ket *na'ñ* ~ *nā'ñ*, Yug *ñe'ñ* id.] (VWJ II 26) < Uralic *ñāñ* ~ *ñāñ* id. << Pers. *nān* id. (Stachowski 1997: 237; Maciuszak 1998: 21; Anikin 2000: 413).

\* \* \*

Some more questions arise when discussing the Yeniseic word for 'beer':

*syrá* (Arin.) ~ *siřá* (Ass.) ~ *sera* || *šera* (Kott: Castrén) ~ *šira* (Kott: Klaproth) ~ *sihirá* (Kott: Mueller) id. = Ottom. *šire* 'juice' ~ Čuv. *sāra* 'beer' ~ Koib. (-Turkic) *sere* id. ~ Oir., Kirgh. *syra* id. < Pers. *šīra* 'juice; syrop, treacle' (ultimately related to English *sour*, Aalto 1971: 30; 198: 7). – Since Pers. *š* always reverts to Old Pers. \**š* (Pisowicz 1985: 174), no older Persian variant like \**šira* has to be assumed. On the other hand, there is no necessity for Turkic to change the initial *š*- > *s*-. Consequently, the existence of *s*- and *š*- variants in Yeniseic cannot be explained as a trace of a diachronic difference in Persian or Turkic. Neither does it look like a regular inner-Yeniseic development which would rather yield *š*- in Ass. and Kott, as in Arin. *sat* ~ Ass., Kott *šet* 'river' (Toporov 1967: 313), Arin. *saj* ~ Ass., Kott *šig* 'night' (VWJ II 206 s.v. *ši*). It points rather to two different channels of borrowing, instead, even if they cannot be established and explained at the moment<sup>4</sup>. – The varying vowel of the first syllable of the Yeniseic words might suggest the Chuvash language as their direct source. On the other hand, Čuv. *sāra* reflects an older form like \**sere*, so that the Yeniseic words should in this case be relatively young loans (but see also below). – G. F. Mueller's spelling <-ihi> in the Kott record certainly is but an orthographical device to render long *-ī-*, so that <sihirá> is to be read *šira*, with the accent on the ultima. – The situation of the final vowel also deserves attention. It is especially conspicuous that none of the Yeniseic words has an *-e* which is typical of almost all Turkic records. It would not be easy to connect this observation with the Persian fluctuation of *a* and *ā* since the word *šira* has no phonetic surrounding in which the alternation can usually be observed<sup>5</sup>. It might be very interesting to suggest that the Yeniseic forms are older than the Turkic ones because they have retained the original Persian pronunciation (*-a*), whereas the Turkic reflexes point to a secondary *-ā* in Persian, but this explanation would not in reality hold its ground. Rather, we have to accept the general Persian pronunciation with *-a* that has, in the Turkic languages, been changed into *-e* according

Aalto's (1987: 104) reservation about this idea is based exactly on the too early attestation (3<sup>rd</sup> century AD) of the word *Čāc*. However, it is not so much philology as rather etymology that makes Pulleyblank's suggestion unacceptable – if, originally, the Yeniseic word really was a nominal compound (Werner 1997) with *-š* as a relict element of the second noun of the syntagma it cannot be used as a one-syllable etymon of *Čāc*. Interestingly enough, H. Werner does not repeat this explanation in VWJ II 312.

<sup>4</sup> Uralic languages should at any rate be excluded as a medium because they display reflexes shortened to one syllable only, like Ostyak *sar* 'beer' and so forth (Aalto 1971: 30).

<sup>5</sup> For the enumeration of these cases see Pisowicz 1985: 15.

to the vowel harmony rules (but cf. also Kirgh. *syra*, in which the first vowel  $\bar{i}$ - > -y-). In short: the (almost regular) distribution of -a and -e in Yeniseic and Turkic cannot provide substantial evidence for chronological or geographical (= time or dialects?) features of the Persian etymon of this word or for establishing the borrowing routes. – Conclusions: The Yeniseic data form three groups with regard to their initial segment: [a] sV-, as in Arin. *syrá*, Ass. *siṛá*, Kott *sera*; [b] sV̄-, as in Kott *sīra* <sihirá>; [c] šV-, as in Kott *šira* ~ *šera*. Incidentally, no Yeniseic word has a segment šV̄-, as in the Persian etymon. – The s variants possibly penetrated through the Turkic medium whereas the borrowing channel of š variants remains unknown since neither a direct borrowing from Pers. *sīra* is certain nor does the Ottoman medium appear realistic.

\* \* \*

The word *kōs* which possibly suggests a relatively early date of the borrowing process seems to be an especially interesting item:

*kōs* (Koib.) ‘beautiful, pretty’ ~ *kos*, in: *Kos-ul’* (Ass.), name of a river, lit. ‘beautiful water’ (VWJ I 458) [= Yug *ku·š* ‘picture’, in: *ejš-ku·š* ‘icon’ (with *ejš* ‘God’) ~ *kūs* ‘1. Gottheit, Geist; 2. anthropomorpher Anhänger an einem Schamanenanzug’ (ibid.); in compounds also: Yug *-ku·s* ~ *-gus*, Southern Ket *-kūs* (ibid.)] < \**kōš* (> Khk. *xos* ‘picture’, Čul. *kōs* [~ *kuas*, see below] ‘beautiful’ < Early New Persian \**x”uš* ~ \**x”oš* ~ \**x”aš* ‘beautiful; nice’ (Stachowski 1996: 98). – For the *x”*- > *x*- development in the 13<sup>th</sup>-14<sup>th</sup> century Persian see Pisowicz 1985: 122. – The Proto-Yeniseic etymon is reconstructed as \**ko’as* or \**kugas* in VWJ I 458. This seems, however, to be due to a misunderstanding. H. Werner (ibid.) tries to find a common etymon for forms like Koib. *kōs* and e.g. Kamassian *kuwas* ‘beautiful’, Čul. (-Turkic) *kuas* id., and he cites my article (Stachowski 1996: 98) by giving the following etymology: «< pers. *kāgaḍ* ‘Papier’», whereas I had in reality suggested a contamination of Pers. *kāgaḍ* ‘paper’ with Pers. *xōš* ‘beautiful; nice’ which has yielded a twofold result in Siberia: [a] the form of *xōš* and the meaning of *kāgaḍ* → Yug *ku·š* ‘picture’; [b] the form of *kāgaḍ* and the meaning of *xōš* → Tuvian *kās* ‘beautiful’. What should be added now is that Koib. *kōs* and Ass. *kos* (which were unknown to me before) are direct (= non-contaminated) reflexes of the Early New Persian word \**x”oš*. Although we have no 18<sup>th</sup> century record of Yug *ku·š*, its form points to an etymon with a long vowel, i.e. an archaic variant \**kōš* which means that it can be viewed as equally old a loan as Koib. *kōs* and Ass. *kos*. They all go back to a period between the 13<sup>th</sup> and 18<sup>th</sup> century.

\* \* \*

Persian sources of the two words below are less certain:

*bešpān* (Ket) ‘spring-balance’ is etymologized from Russ. *безмѣн* id. in VWJ I 122. However, this leaves the change of the palatalness of Russ. -ѣ- > Ket -ā- unexplained. I would rather suggest an etymon like \**besmān* (> Old Čuv. \**besmen* [> modern Čuv. *pasman*, a measure of weight of cereals] > Russ. *безмѣн* ‘spring-balance’) < Pers. Pl. *vaznān* id. < Sg. \**vazn* < Arab. *wazn* ‘weight (a piece of metal)’; semantically cf. Russ. *вѣс* ‘weight’ vs. *вѣсъ* ‘balance’. – For the time being, I cannot establish the language with which the etymon \**besmān* should be connected. – Besides, there is no hint about the chronology of the borrowing of this word into Ket.

*etnā* (Imb.) «‘leicht’ (fehlerhaft anstatt ‘lebendig?’)» (VWJ I 253) might be somehow connected with Pers. *āsān* ‘light’ >> Oir. *āzān* ‘healthy’, Barab. *esen* ~ *izen* id., Khak. *izen* id., Kirgh. *esen* ‘1. happy; 2. healthy’ (Stachowski 1993: 249). – The

change of Pers. *-s-* > Imb. *-t-* could have actually been caused by a folk-etymological association of the loan word with Ket *et* 'alive, living' but one question still remains unanswered: why did Pers. *-ān* change into Imb. *-na*?

## 2.

The approximate geographical distribution of the languages important to us in this context was in the 18<sup>th</sup> century as follows:

Ostyak	Selkup	
	Ket	
Pump.		
	Arin.	
		Kott
		Ass.
	Kamassian	
Shor		
Khak.	Mator	Tof.
Oir.		Tuv.

Now, let us try to establish the main channels of borrowing of the Persian words presented above (apart from *bešpān* and *emá* which are uncertain).

The word *nañ* 'bread' is unknown to the Turkic languages of Siberia, so that it should be seen as a Uralic loan into Yeniseic. Ostyak and Selkup seem to be the only Uralic languages that should be taken into account for that purpose. At the same time, *nañ* is the only Persian word that can be assumed to have come from the North, here-with exemplifying the ties between the languages of an areal grouping called by E. Helimski (2003: 160) an "Ostyak (Ob-Yeniseic) Sprachbund".

It seems especially alluring to see whether other conceivable borrowing routes of the Persian words can also be correlated with some areal groupings presented in the study by Helimski. The Persian loans are attested, besides Yeniseic, in the following languages of the region:

*ajna* – Shor, Khakas, Oirot

*molat* – Khakas; *bolat* – Mator, Tofalar, Mongolian

*syra* – Oirot

*kōs* – Khakas.

The first thing to be observed here is that *molat* and *bolat* can be interpreted as two phonetic variants which penetrated via two different routes into Yeniseic. The routes can, on the geographical basis, be called a Western (Khakas) and an Eastern (Tofalar) one<sup>6</sup>. Moreover, what is even more important is that there probably exists not a single Persian word which was transmitted through the Eastern (Tuvinian, Tofalar) channel only. Both *molat* and the three remaining words (*ajna*, *syra*, *kōs*) point to the Western route (Oirot, Khakas, Shor) as the main channel of borrowing (even if Khak. *xos* 'picture' cannot be accepted as a direct source of Yen. *kōs* 'beautiful' [see the discussion above], not a vaguest hint about the Eastern borrowing route of this word exists).

<sup>6</sup> The Mongolian mediation can, as it seems, be ignored here because there exists no other Persian loan word transmitted through Mongolian into Yeniseic, and no peculiar feature of this word compels us to abandon the prospect of the Turkic mediation.

The limited number of Persian words in the Yeniseic languages of the 18<sup>th</sup> century does not permit far-reaching inferences. We may, however, glean from what has been said above that this lexical material points to the participation of Kott and Arinian in the "Upper Yenisei Sprachbund" (Helimski 2003: 158) whose core languages are Khakas, Shor, Kamassian and Mator, whereas the cautious conjecture that Tuvian and Tofalar, too, may belong to this Sprachbund has not found any support in my material. This fact makes the real existence of a grouping, supplied by Helimski with a question mark, viz. the "(?) Yeniseic Sprachbund" (ibid. 161) which, besides Yeniseic, comprises also Tuvian and Tofalar even more uncertain. The glottal articulation of some vowels is most apparently not sufficient to suggest a discrete linguistic area so that I willingly agree with E. Helimski's suggestion concerning this phonetic phenomenon: "[...] it could be safer to assume instead a chain of contact-induced developments, with their starting point presumably in Samoyedic" (ibid.) and I am even inclined to go a step further and to assume a mere transmission of this articulatory habit by Yeniseic-speakers in the course of their Turkicization which in fact is a contact-induced development, but one reduced to a stretch between two nodes only instead of a chain of developments.

## REFERENCES

- NVS = Mueller, G. F.: *Nachrichten über Völker Sibiriens (1736-1742)*, ed. by E. Helimski & (+) H. Katz, Hamburg 2003.
- ŞS = Kurpeško-Tannagaşeva, N. N. / Akalın, Ş. H.: *Şor sözlüğü*, Adana 1995.
- VWJ = Werner, H.: *Vergleichendes Wörterbuch der Jenissej-Sprachen*, vol. 1-3, Wiesbaden 2002.
- Aalto, P. 1971: Iranian contacts of the Turks in pre-islamic times. – Ligeti, L. (ed.): *Studia Turcica*, Budapest: 29-37.
- 1987: The name of Tashkent. – [Aalto, P.:] *Studies in Altaic and comparative philology. A collection of Professor Pentti Aalto's essays in honour of his 70<sup>th</sup> birthday* (= *Studia Orientalia* 59), Helsinki: 103-108 [reprint from: *Central Asiatic Journal* 21 (1977): 193-198].
- 1998: Finnish *olut* and *humala*, once again. – *Oekeeta asijoo. Commentationes Fenno-Ugricae in honorem Seppo Suhonen sexagenarii* (= *MSFOu* 228), Helsinki: 1-11.
- Anikin, A. E. 2000: *Этимологический словарь русских диалектов Сибири. Заимствования из уральских, алтайских и палеоазиатских языков*, <sup>2</sup>Novosibirsk.
- Helimski (= Хелимский), E. A. 1986: Архивные материалы XVIII века по енисейским языкам. – Сборник *Палеоазиатские языки*, Leningrad: 179-213.
- 1997: *Die matorische Sprache. Wörterverzeichnis – Grundzüge der Grammatik – Sprachgeschichte*, Szeged.
- 2003: Areal groupings (Sprachbünde) within and across the borders of the Uralic language family: A survey. – *Nyelvtudományi Közlemények* 100: 156-167.
- Joki, A. J. 1952: *Die Lehnwörter des Sajansamojedischen*, Helsinki.
- Maciuszak, K. 1998: New Persian *nān* 'bread' and its reflexes in Siberian languages. – *Studia Etymologica Cracoviensia* 3: 17-23.
- Munkácsi, B. 1900: Todesdæmon *Aina* der Altaier = Todesdæmon *Xiñ* der Wogulen. – *Keleti Szemle* 1: 158-160.
- Pisowicz, A. 1985: *Origins of the New and Middle Persian phonological systems*, Kraków.
- Stachowski, M. 1993: Beiträge zur Kenntnis der arabischen und persischen Lehnwörter in den südsibirischen Türksprachen. – *Folia Orientalia* 29: 247-259.

- 
- 1996: Über einige altaische Lehnwörter in den Jenissej-Sprachen. – *Studia Etymologica Cracoviensia* 1: 91-115.
- 1997: Altaistische Anmerkungen zum “Vergleichenden Wörterbuch der Jenissej-Sprachen”. – *Studia Etymologica Cracoviensia* 2: 227-239.
- 2002: Заимствованные слова в долганском языке (Обзор проблем). – *Kyoto University Linguistic Research* 21: 1-24.
- Торогов, V. N. 1967: Из этимологии енисейских языков (К вопросу об одном ряде соответствий пумпокольскому *t*). – *Этимология 1965* (ed. 1967): 311-320.
- Vovin, A. 2000: Did the Xiong-nu speak a Yeniseian language? – *Central Asiatic Journal* 44/1: 87-104.
- 2003: Did the Xiong-nu speak a Yeniseian language? Part 2: Vocabulary. – Sárközi, A. / Rákos, A. (ed.): *Altaica Budapestinensia MMII*, Budapest: 389-394.
- Werner, H. 1994: *Das Klassensystem in den Jenissej-Sprachen*, Wiesbaden.
- 1997: Das jensejische und türkische Wort für ‘Stein’ im Lichte der Pluralbildung der jensejischen Nomina. – *Studia Etymologica Cracoviensia* 2: 275-282.